CHAPTER 3

Method—Not Trial-and-Error

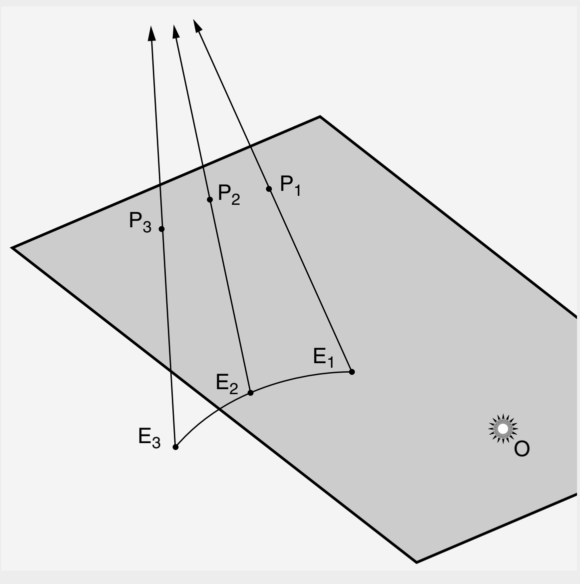
In investigations such as we are now pursuing, it should not be so much asked "what has occurred," as "what has occurred that has never occurred before."

—**C. Auguste Dupin,**
in Edgar Allan Poe's
"The Murders in the Rue Morgue"

With Dupin's words in mind, let us return to the dilemma in which we had entangled ourselves in our discussion in the previous chapter. That dilemma was connected to the fact, that what Piazzi observed as the motion of the unknown object against the fixed stars, was neither the object's actual path in space, nor even a simple projection of that path onto the celestial sphere of the observer, but rather, the result of the motion of the object and the motion of the Earth, mixed together.

Thanks to the efforts of Kepler and his followers, the determination of the orbit of the Earth, subsuming its distance and position relative to the sun on any given day of the year, was quite precisely known by Gauss's time. Accordingly, we can formulate the challenge posed by Piazzi's observations in the following way: We can determine a precise set of positions in space from which Piazzi's observations were made, taking into account the Earth's own motion. From each of the positions of Palermo, where Piazzi's observatory was located, draw a straight line-of-sight in the direction in which Piazzi saw the object at that moment. All we can say with certainty about the actual positions of the unknown object at the given times, is that each position lies *somewhere* along the corresponding straight line. What shall we do?

FIGURE 3.1. *Piazzi's observations define three "lines of sight" from three Earth positions E_1, E_2, E_3, but do not tell us where the planet lies on any of those lines. We do know that the positions lie on some plane through the sun.*



In the face of such an apparent degree of ambiguity, any attempt to "curve fit" fails. For, there are no well-defined positions on which to "fit" an orbit! But, don't we know *something* more, which could help us? After all, Kepler taught that the geometrical *forms* of the orbits are (to within a very high degree of precision, at least) plane conic sections, having a common focus at the center of the sun. Kepler also provided a crucial, additional set of constraints (to be examined in Chapter 7), which determine the precise motion in any given orbit, once the "elements" of the orbit discussed last chapter have been determined.

Now, unfortunately, Piazzi's observations don't even tell us what *plane* the orbit of Piazzi's object lies in. How do we find the right one?

Take an arbitrary plane through the sun. The lines-of-sight of Piazzi's observations will intersect that plane in as many points, each of which is a candidate for the position of the object at the given time. Next, try to construct a conic section, with a focus at the sun, which goes through those points or at least fits them as closely as possible. (Alas! We are back to curve-fitting!) **(Figure 3.1)**

Finally—and this is the substantial new feature—check whether the time intervals defined by a Keplerian motion along the hypothesized conic section between the given points, agree with the actual time intervals of Piazzi's observations. If they don't fit, which will be nearly always the case, then we reject the orbit. For example, if the intersection-points are very far away from the sun, then Kepler's constraints would imply a very slow motion in the corresponding orbit; outside a certain distance, the corresponding time-intervals would become larger than the times between Piazzi's actual observations. Conversely, if the points are very close to the sun, the motion would be too fast to agree with Piazzi's times.

The consideration of time-intervals thus helps to limit the range of trial-and-error search somewhat, but the domain of apparent possibilities still remains monstrously large. With the unique exception of Gauss, astronomers

C.F. Gauss: 'To determine the orbit of a heavenly body, without any hypothetical assumption'

It seems somewhat strange that the general problem—*to determine the orbit of a heavenly body, without any hypothetical assumption, from observations not embracing a great period of time, and not allowing a selection with a view to the application of special methods*—was almost wholly neglected up to the beginning of the present century; or, at least, not treated by any one in a manner worthy of its importance; since it assuredly commended itself to mathematicians by its difficulty and elegance, even if its great utility in practice were not apparent. An opinion had universally prevailed that a complete determination from observations embracing a short interval of time was impossible,—an ill-founded opinion,—for it is now clearly shown that the orbit of a heavenly body may be determined quite nearly from good observations embracing only a few days; and this without any hypothetical assumption.

Some ideas occurred to me in the month of September of the year 1801, [as I was] engaged at that time on a very different subject, which seemed to point to the solution of the great problem of which I have spoken. Under such circumstances we not infrequently, for fear of being too much led away by an attractive investigation, suffer the associations of ideas, which, more attentively considered, might have proved most fruitful in results, to be lost from neglect. And the same fate might have befallen these conceptions, had they not happily occurred at the most propitious moment for their preservation and encouragement that could have been selected. For just about this time the report of the new planet, discovered on the first day of January of that year with the telescope at Palermo, was the subject of universal conversation; and soon afterwards the observations made by that distinguished astronomer Piazzi, from the above date to the eleventh of February were published.

Nowhere in the annals of astronomy do we meet with so great an opportunity, and a greater one could hardly be imagined, for showing most strikingly, the value of this problem, than in this crisis and urgent necessity, when all hope of discovering in the heavens this planetary atom, among innumerable small stars after the lapse of nearly a year, rested solely upon a sufficiently approximate knowledge of its orbit to be based upon these very few observations. Could I ever have found a more seasonable opportunity to test the practical value of my conceptions, than now in employing them for the determination of the orbit of the planet Ceres, which during these forty-one days had described a geocentric arc of only three degrees, and after the lapse of a year must be looked for in a region of the heavens very remote from that in which it was last seen?

The first application of the method was made in the month of October 1801, and the first clear night (December 7, 1801), when the planet was sought for as directed by the numbers deduced from it, restored the fugitive to observation. Three other new planets subsequently discovered, furnished new opportunities for examining and verifying the efficiency and generality of the method. [*emphasis in original*]

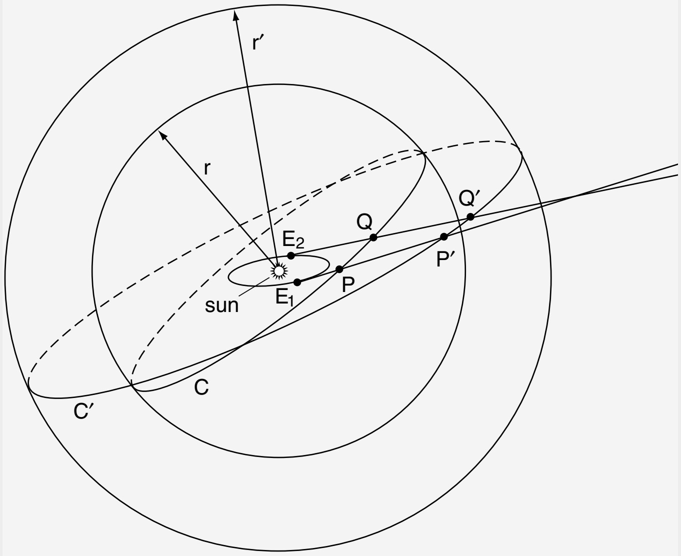
Excerpted from the Preface to the English edition of Gauss's "Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections."

felt themselves forced to make ad hoc assumptions and guesses, in order to radically reduce the range of possibilities, and thereby reduce the trial-and-error procedures to a minimum.

For example, the astronomer Wilhelm Olbers and others decided to start with the working assumption that the sought-for orbit was very nearly circular, in which case the motion becomes particularly simple. Kepler's third constraint (usually referred to as his "Third Law") determines a specific rate of uniform motion along the circle, as soon as the radius of the circular orbit is known. According to that third constraint, the square of periodic time in any closed orbit—i.e., a circular or an elliptical one—as measured in years, is equal to the cube of the orbit's major axis, as measured in units of the major axis of the Earth's orbit. Next, Olbers took **two** of Piazzi's observations, and calculated the radius which a circular orbit would have to have, in order to fit those two observations.

It is easy to see how to do that in principle: The two observations define two lines of sight, each originating from the position of the Earth at the moment of observation. Imagine a sphere of variable radius r, centered at the sun. **(Figure 3.2)** For each choice of r, that sphere will intersect the lines-of-sight in two points, P and Q. Assuming the planet were actually moving on a circular orbit of radius r, the points P and Q would be the corresponding positions at the times of the two observations, and the orbit would be the great circle on the sphere passing through those two points. On the other hand, Kepler's constraints tell us exactly how large is the arc which any planet would traverse, during the time interval between the two observations, if its orbit were a circle of radius r. Now compare the arc determined from

FIGURE 3.2. *Method to determine the orbit of Ceres, on the assumption that the orbit is circular. Two sightings of Ceres define two lines of sight coming from the Earth positions E_1, E_2 (the Earth's positions at the moments of observation). A sphere around the sun, of radius r, intersects the lines of sight in two points P,Q, which lie on a unique great circle C on that sphere. A sphere of some different radius r′ would define a different set of points P′, Q′ and a different hypothetical orbit C′. Determine the unique value of r, for which the size of the arc PQ agrees with the rate of motion a planet would really have, if it were moving according to Kepler's laws on the circular orbit C over the time interval between the given observations.*



Kepler's constraint, with the actual arc between *P* and *Q*, as the length of radius *r* varies, and locate the value or values of *r*, for which the two become coincident. That determination can easily be translated into a mathematical equation whose numerical solution is not difficult to work out. Having found a circular orbit fitting two observations in that way, Olbers then used the comparison with other observations to correct the original orbit.

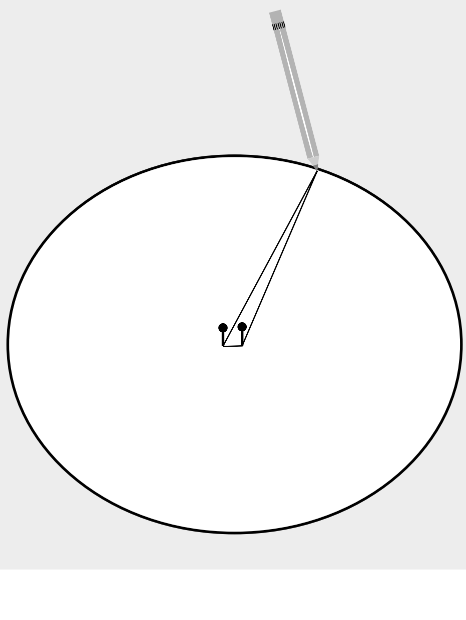
Toward the end of 1801 astronomers all over Europe began to search for the object Piazzi had seen in January-February, based on approximations such as Olbers'. The search was in vain! In December of that year, Gauss published his hypothesis for the orbit of Ceres, based on his own, entirely new method of calculation. According to calculations based on Gauss's elements, the object would be located more than 6° to the south of the positions forecast by Olbers, an enormous angle in astronomical terms. Shortly thereafter, the object was found very close to the position predicted by Gauss.

Characteristically, *Gauss's method used no trial-and-error at all.* Without making any assumptions on the particular form of the orbit, and using only three well-chosen observations, Gauss was able to construct a good first approximation to the orbit immediately, and then perfect it *without further observations* to a high precision, making possible the rediscovery of Piazzi's object.

To accomplish this, Gauss treated the set of observations (including the times as well as the apparent positions) as being the equivalent of a *set of harmonic intervals.* Even though the observations are, as it were, jumbled up by the effects of projection along lines-of-sight and motion of the Earth, we must start from the standpoint that the underlying curvature, determining an entire orbit from any arbitrarily small segment, is somehow lawfully expressed in such an array of intervals. To determine the orbit of Piazzi's object, we must be able to identify the specific, tell-tale characteristics which reveal the whole orbit from, so to speak, "between the intervals" of the observations, and distinguish it from all other orbits. This requires that we conceptualize the higher curvature underlying the entire manifold of Keplerian orbits, taken as a whole. Actually, the higher curvature required, cannot be adequately expressed by the sorts of mathematical functions that existed prior to Gauss's work.

We can shed some light on these matters, by the following elementary experimental-geometrical investigation. Using the familiar nails-and-thread method, con-

FIGURE 3.3. *Constructing an ellipse in the shape of the orbit of Mars.*



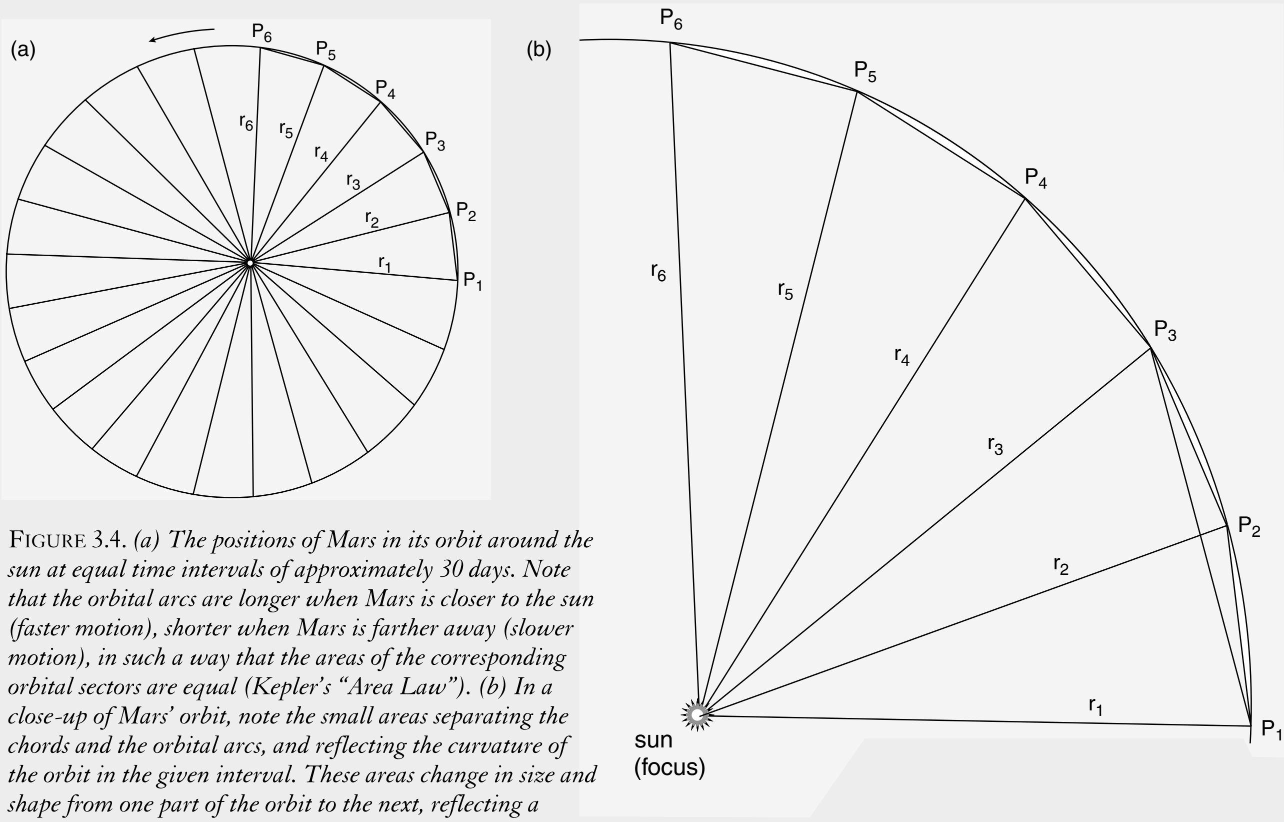


FIGURE 3.4. *(a) The positions of Mars in its orbit around the sun at equal time intervals of approximately 30 days. Note that the orbital arcs are longer when Mars is closer to the sun (faster motion), shorter when Mars is farther away (slower motion), in such a way that the areas of the corresponding orbital sectors are equal (Kepler's "Area Law"). (b) In a close-up of Mars' orbit, note the small areas separating the chords and the orbital arcs, and reflecting the curvature of the orbit in the given interval. These areas change in size and shape from one part of the orbit to the next, reflecting a constantly changing curvature.*

struct an ellipse having the shape of the Mars orbit, as follows. **(Figure 3.3)** Hammer two nails into a flat board covered with white paper, at a distance of 5.6 cm from each other. Take a piece of string 60 cm long and tie each end to one of the nails—or alternatively, make a loop of string of length 60 + 5.6 = 65.6 cm, and loop it around both nails. Pulling the loop tight with the tip of a pencil as shown, trace an ellipse. The positions of the two nails represent the foci. The resulting curve will be a scaled-down replica of Mars' orbit, with the sun at one of the foci.

Observe that the circumference generated is hardly distinguishable, by the naked eye, from a circle. Indeed, mark the midpoint of the ellipse (which will be the point midway between the foci), and compare the distances from various points on the circumference, to the *center.* You will find a maximum discrepancy of only about one millimeter (more precisely, 1.3 mm), between the maximum distance (the distance between the points on the circumference at the two ends of the major axis connecting the two foci) and the minimum distance (between the endpoints of the minor axis drawn perpendicular to the major axis at its mid-point). Thus, this ellipse's deviation from a perfect circle is only on the order of four parts in one thousand. How was Kepler able to detect and demonstrate the non-circular shape of the orbit of Mars, given such a minute deviation, and how could he correctly ascertain the precise nature of the non-circular form, on the basis of the technology available at his time?

Observe in **Figure 3.4a**, that the distances to the *sun* (the marked focus) change *very substantially***,** as we move along the ellipse.

Now, choose two points P_1 and P_2 anywhere along the circumference of the ellipse, two centimeters apart. The interval between them would correspond to successive positions of Mars at times about seven days apart (actually, up to about 10 percent more or less than that, depending on exactly where P_1 and P_2 lie, relative to the *perihelion* [closest] and *aphelion* [farthest] positions). Draw radial lines from each of P_1, P_2 to the sun, and label the corresponding lengths r_1, r_2.

Consider what is contained in the *curvilinear triangle* formed by those two radial line segments and the small arc of Mars' trajectory, from P_1 to P_2. Compare that arc with that of analogous arcs at other positions on the orbit, and consider the following propositions: Apart from the symmetrical positions relative to the two axes of the ellipse, *no two such arcs are exactly superimposable in any of their parts.* Were we to change the parameters of the ellipse—for example, by changing the distance between the foci, by any amount, however small—then *none* of the arcs on the new ellipse, no matter how small, would be superimposable with *any* of those on the first, in any of their parts! Thus, each arc is uniquely characteristic of the ellipse of which it is a part. The same is true among all species of Keplerian orbits.

Consider what means might be devised to reconstruct the whole orbit from any one such arc. For example, by what means might one determine, from a small portion of a planetary trajectory, whether it belongs to a parabolic, hyperbolic, or elliptical orbit?

Now, compare the orbital arc between P_1 and P_2 with the straight line joining P_1 and P_2. **(Figure 3.4b)** Together they bound a tiny, virtually infinitesimal area. Evidently, the unique characteristic of the particular elliptical orbit must be reflected somehow in the *specific manner* in which that arc *differs* from the line, as reflected in that "infinitesimal" area.

Finally, add a third point, P_3, and consider the curvilinear triangles corresponding to each of the three pairs (P_1,P_2), (P_2,P_3), and (P_1,P_3), together with the corresponding rectilinear triangles and "infinitesimal" areas which compose them. The harmonic mutual relations among these and the corresponding time intervals, lie at the heart of Gauss's method, which is *exactly the opposite* of "linearity in the small."

—JT

Chapter 4

Families of Catenaries

(An Interlude Considering Some Unexpected Facts About 'Curvature')

Any successful solution of the problem posed to Gauss must pivot on conceptualizing the characteristic curvature of Keplerian orbits "in the small." Before turning to Kepler's own investigations on this subject, it may be helpful to take a brief look at the closely related case of families of catenaries on the surface of the Earth—these being more easily accessible to direct experimentation, than the planetary orbits themselves.

Catenaries, Monads, and A First Glimpse at Modular Functions

When a flexible chain is suspended from two points, and permitted to assume its natural form under the action of its own weight, then, the portion of the chain between the two points forms a characteristic species of curve, known as a catenary. The ideal catenary is generated by a chain consisting of very small, but strong links made of a rigid material, and having very little friction; such a chain is practically inelastic (i.e., does not stretch), while at the same time being nearly perfectly flexible, down to the lower limit defined by the diameter of the individual links.

Interestingly, the form of the catenary depends only on the position of the points of suspension and the length of the chain between those points, but not on its mass or weight.

With the help of a suitable, fine-link chain, suspended parallel to, and not far from, a vertical wall or board (so that the chain's form can easily be seen and traced, as desired), carry out the following investigations.

(For some of these experiments, it is most convenient to use two nails or long pins, temporarily fixed into the wall or board, as suspension-points; the nails or pins should be relatively thin, and with narrow heads, so that the links of the chain can easily slip over them, in order to be able to vary the length of the suspended portion. In some experiments it is better to fix only one suspension-point with a nail, and to hold the other end in your hand.)

Start by fixing any two suspension-points and an arbitrary chain-length. **(Figure 4.1)** Observe the way the shape of each part of the catenary, so formed, depends on all the other parts. Thus, if we try to modify any portion of the catenary, by pushing it sideways or upwards with

Figure 4.1 *A catenary is formed by suspending a chain between points A and B.*

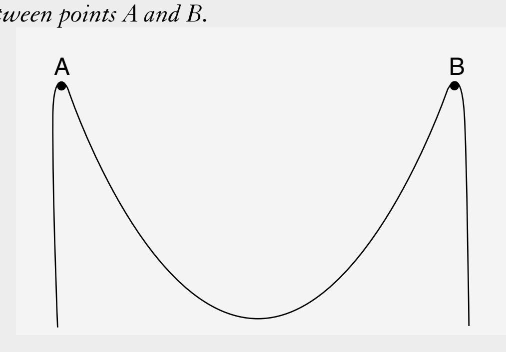
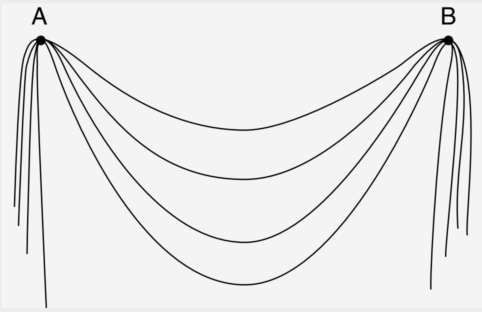


FIGURE 4.2. *Varying the lengths of the chain generates a family of catenaries of varying curvatures.*



the tip of a finger, we see that the entire curve is affected, at least slightly, over its entire length. This behavior of the catenary reflects Leibniz's principle of least action, whereby the entire Universe as a whole, including its most remote parts, reacts to any event anywhere in the Universe. There is no "isolated" point-to-point action in the way the Newtonians claim.

Note that the curvature of each individual catenary changes constantly along its length, as we go from its lowest point to its highest point.

Next, generate a family of catenaries, by keeping the suspension-points fixed, but varying the length of the chain between those points. **(Figure 4.2)** Observe the changes in the form and curvature, and the changes in the angles, which the chain makes to the horizontal at the points of suspension, as a function of the suspended length.

Generate a second family of catenaries, by keeping the chain length and one of the suspension-points fixed, while varying the other point. **(Figure 4.3)** If *A* is the first suspension-point, and *L* is the length of the suspended chain, then the second suspension-point *B* (preferably held by hand) can be located anywhere within the circle of radius *L* around *A*. For *B* on the circumference of the circle, the catenary degenerates into a straight line. (Or rather, something close to a straight line, since the latter would require a physically impossible, "infinite tension" to overcome the gravitational effect.) Observe the changes of form, as *B* moves around *A* in a circle of radius less than *L*. Also, observe the change in the angles, which the catenary makes to the horizontal at each of the endpoints, as a function of the position of *B*. Finally, observe the changes in the tension, which the chain exerts at the endpoint *B*, held by hand, as its position is changed.

Examine this second family of catenaries for the case, where the suspended length is extremely short. Combining the variation of the endpoint with variation of length

FIGURE 4.3. *Varying the endpoint position of a fixed length of chain generates a second family of catenaries.*

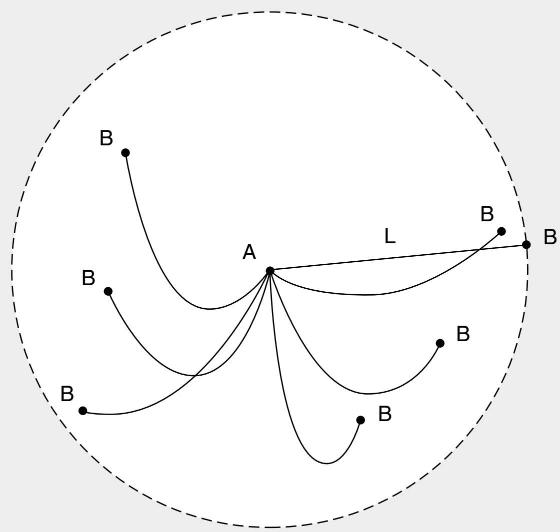
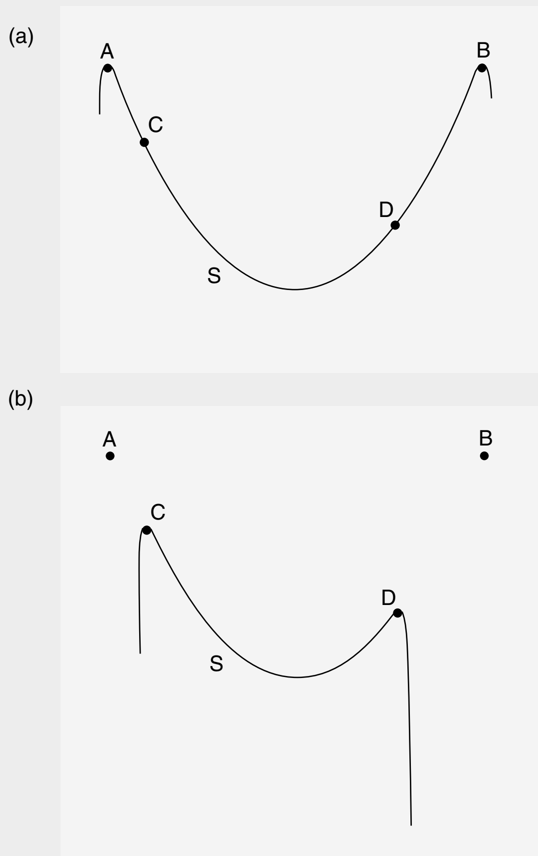


FIGURE 4.4. *Release catenary AB to points C,D. Every arc of a catenary, is itself a catenary!*



(families one and two) gives us the manifold of all elementary catenaries.

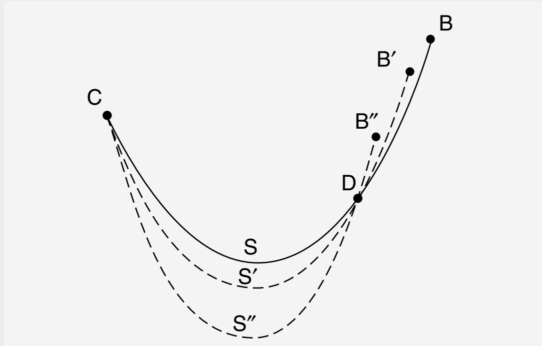
Consider, next, the following remarkable proposition: *Every arc of a catenary, is itself a catenary!* To wit: On a catenary with fixed suspension-points *A,B,* examine the arc *S* bounded by any two points *C* and *D* on the curve. (**Figure 4.4a**) Drive nails through the chain at *C* and *D* into the wall or board behind it. Note that the form of the chain remains unchanged. If we then remove the parts of the chain on either side of the arc, or simply release the chain from its original supports *A* and *B*, then the portion of the chain between *C* and *D* will be suspended from those points as a catenary, while still retaining the original form of the arc *S*. **(Figure 4.4b)**

Consider another remarkable proposition: *The entire form of a catenary (up to its suspension-points), is implicitly determined by any of its arcs, however small.* Or, to put it another way: If any arc of one catenary, however small, is congruent in size and shape to an arc on another catenary, then the two catenaries are superimposable over their *entire* lengths. (Only the endpoints might differ, as when we replaced *A,B* by *C,D* to obtain a subcatenary of an originally longer catenary.) To get some insight into the validity of this proposition, try to "beat" it by an experiment, as follows.

Fix one of the endpoints of the arc in question, say *C,* by a nail, and mark the position of the other endpoint, *D,* on the wall or board behind the chain. (**Figure 4.5**) Now taking the end of the chain on *D*'s side, say *B,* in your hand (i.e., the right-hand endpoint, if *D* is to the right of *C,* or *vice versa*), try to move that endpoint in such a way, that the corresponding catenary, whose other suspension-point is now *C,* always passes through the position *D* as verified by the mark on the adjacent wall or board. Holding to that constraint, we generate a family of catenaries having the two common points *C* and *D*. In doing so, observe that the shape of the arc between *C* and *D* continually changes, as the position of the movable endpoint *B* is changed. This change in shape correlates with the observation, that the tension exerted by the chain at its endpoints, changes according to their relative positions; according to the higher or lower level of tension, the arc between *C* and *D* will be less or more curved. Only a single, unique position of *B* (namely, the original one) produces exactly the same tension and same curvature, as the original arc *CD.* Our attempt to "beat" the stated proposition, fails.

While admittedly deserving more careful examination, these considerations suggest three things: Firstly, that all the catenary arcs, which are parts of one and the same catenary, share a common internal characteristic, which in turn determines the larger catenary as a whole. In consequence of this, secondly, when we look at different parts of a given catenary, *we are in a sense looking at different local expressions on the same global entity.* Although various, small portions of the catenary have different curvatures in the sense of visual geometry, in a deeper sense they all share a common "higher curvature," characteristic of the catenary of which they are parts. Finally, there must be a still higher mode of curvature, which defines the common characteristic of the entire family of catenaries. That latter entity would be congruent with Gauss's concept of a modular function for the species of catenaries, as a special case of his hypergeometric function; the latter subsuming the catenaries together with the analogous, crucial features of the Keplerian planetary orbits. (In the Earth-bound case of elementary catenaries, the distinction among different catenaries is, to a very high degree of approximation, merely one of self-similar "scaling." That is not even approximately the case for Keplerian orbits.)

FIGURE 4.5. *Only one unique position of B produces the exact tension and curvature of catenary CD. Different parts of a given catenary are local expressions of the whole, sharing a common internal characteristic.*



In a 1691 paper on the catenary problem, Leibniz notes that Galileo had made the error of identifying the catenary with a parabola. Galileo's error, and the discrepancy between the two curves, was demonstrated by Joachim Jungius (1585-1657) through careful, direct experiments. However, Jungius did not identify the true law underlying the catenary. Leibniz stressed, that the catenary cannot be understood in terms of the geometry we associate with Euclid, or, later, Descartes, but *is* susceptible to a *higher form of geometrical analysis,* whose principles are embodied in the so-called "infinitesimal calculus." The latter, in turn, is Leibniz's answer to the challenge, which Kepler threw out to the world's geometers in his *New Astronomy (Astronomia Nova)* of 1609.

—*JT*

Chapter 5

Kepler Calls for a 'New Geometry'

Non-linear curvature, exemplified by our exploration of catenaries, stands in the forefront of Johannes Kepler's revolutionary work *New Astronomy.* There Kepler bursts through the limitations of the Copernican heliocentric model, where the planetary orbits were assumed *a priori* to be circular.

The central paradox left by Aristarchus and Copernicus was this: Assume the motions of the planets as seen from the Earth—including the bizarre phenomena of retrograde motion—are due to the fact that the Earth is not stationary, but is itself moving in some orbit around the sun. These apparent motions result from combinations of the *unknown* true motion of the Earth and the *unknown* true motion of the heavenly bodies. How can we determine the one, without first knowing the other?

In the *New Astronomy,* Kepler recounts the exciting story, of how he was able to solve this paradox by a process of "nested triangulations," using the orbits of Mars and the Earth. Having finally determined the precise motions of *both,* a new set of anomalies arose, leading Kepler to his astonishing discovery of the elliptical orbits and the "area law" for non-uniform motion. Kepler's breakthrough is key to Gauss's whole approach to the Ceres problem, one hundred fifty years later. It is therefore fitting that we examine certain of Kepler's key steps in this and the following chapter.

As to mere shape, in fact, the orbits of the Earth, Mars, and most of the other planets (with the exception of Mercury and Pluto) are very nearly perfect circles, deviating from a perfect circular form only by a few parts in a thousand. The centers of these near-circles, on the other hand, do not coincide with the sun! Consequently, there is a constant variation in the distance between the planet and the sun in the course of an orbit, ranging between the extreme values attained at the *perihelion* (shortest distance) and the *aphelion* (farthest distance).

As Kepler noted, the perihelion and aphelion are at the same time the chief singularities of change in the planet's rate of motion along the orbit: the maximum of velocity occurs at the perihelion, and the minimum at the aphelion.

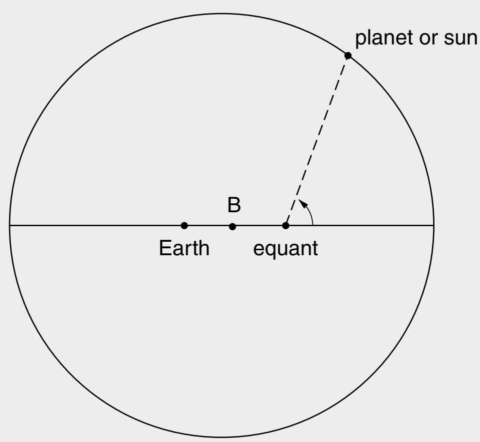
In an attempt to account for this fact, while trying to salvage the hypothesis of simple circular motion as elementary, Ptolemy had devised his theory of the "equant." According to that theory, the Earth is no longer the exact center of the motion, but rather another point *B*. **(Figure 5.1)** The planet is "driven" around its circular orbit (called an "eccentric" because of the displacement of its center from the position of the Earth) in such a way, that its *angular* motion is uniform with respect to a third point (the "equant"), located on the line of apsides opposite the Earth from the center of the eccentric circle.* In other words, the planet moves as if it were swept along the orbit by a gigantic arm, pivoted at the equant and turning around it at a constant rate.

On the basis of his precise data for Earth and Mars, Kepler was able to demolish Ptolemy's equant once and for all. This immediately raised the question: If simple rotational action is excluded as the underlying basis for planetary motion, then what new principle of action should replace it?

Step-by-step, already beginning in the *Mysterium Cosmographicum (Cosmographic Mystery),* Kepler developed his "electromagnetic" conception of the solar system, referring directly to the work of the English scientist William Gilbert, and implicitly to the investigations of Leonardo da Vinci and others on light, as well as Nicolaus of Cusa. Kepler identifies the sun as the original

* Readers should remember that in Ptolemy's astronomical model, the sun and planets are supposed to orbit about the Earth.

Figure 5.1. *To account for the differing rates of motion of the planet, Ptolemy's description placed the Earth at an eccentric (off-center) location, with the planet's uniform angular motion centered at a third, "equant" point.*



source and "organizing center" of the whole system, which is "run" on the basis of a harmonically ordered, but otherwise *constantly changing activity* of the sun vis-à-vis the planets. Kepler's conception of that activity, has nothing to do with the axiomatic assumption of smooth, featureless, linear forms of "push-pull" displacement in empty space, promoted by Sarpi and Galileo, and revived once more in Newton's solar theory, in which the sun is degraded to a mere "attracting center."

On the contrary! According to Kepler, the solar activity generates a harmonically ordered, everywhere-dense array of *events of change,* whose *ongoing, cumulative result* is reflected in—among other things—the visible motion of the planets in their orbits.

The need to elaborate a new species of mathematics, able to account for the *integration* of dense singularities, emerges ever more urgently in the course of the *New Astronomy,* as Kepler investigates the revolutionary implications of his own observation, that *the rate of motion of a planet in its orbit is governed by its distance from the sun.* This relationship emerged most clearly, in comparing the motions at the perihelion and aphelion. The ratio of the corresponding velocities was found to be precisely equal to the *inverse* ratio of the two extreme radial distances. For good reasons, Kepler chose to express this, not in terms of *velocities,* but rather in terms of the *time* required for the planet to traverse a given section of its orbit.*

Kepler's Struggle with Paradox

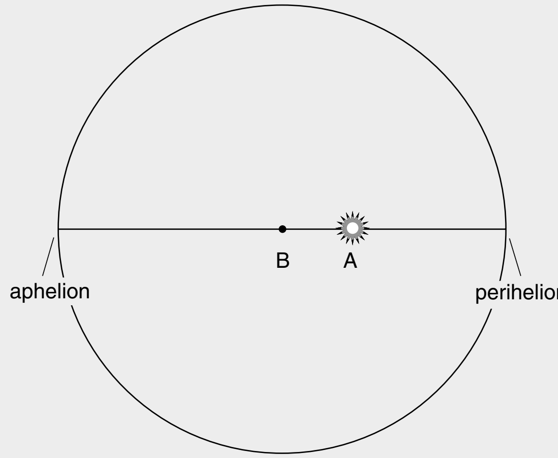
Let us join Kepler in his train of thought. While still operating with the approximation of a planetary orbit as an "eccentric circle," Kepler formulates this relationship in a preliminary way as follows: It has been demonstrated,

> that the elapsed times of a planet on equal parts of the eccentric circle (or equal distances in the ethereal air) are in the same ratio as the distances of those spaces from the point whence the eccentricity is reckoned [i.e., the sun–JT]; or more simply, to the extent that a planet is farther from the point which is taken as the center of the world, it is less strongly urged to move about that point.

Since the distances are constantly *changing,* the existence of such a relationship immediately raises the question: How does the temporally extended motion—as, for example, the periodic time corresponding to an entire revolution of the planet—relate to the magnitudes of those constantly varying "urges" or "impulses"?

* *Cf.* Fermat's later work on least-time in the propagation of light.

FIGURE 5.2. *Kepler's original hypothesis: The planetary orbits are circles whose centers are somewhat eccentric with regard to the sun. Kepler observed that the planet moves fastest at the perihelion, slowest at the aphelion, in apparent inverse proportion to the radial distances.*



A bit later, Kepler picks up the problem again. To follow Kepler's discussion, draw the following diagram. **(Figure 5.2)** Construct a circle and its diameter and label the center *B.* To the right of *B* mark another point *A.* The circumference of the circle represents the planetary orbit, and point *A* represents the position of the sun. Kepler writes:

> Since, therefore, the times of a planet over equal parts of the eccentric, are to one another, as the radial distances of those parts [from the sun–JT], and since the individual points of the entire . . . eccentric are all at different distances, it was no easy task I set myself, when I sought to find how one might obtain the *sums* of the individual radial distances. For, unless we can find the sum of all of them (and they are infinite in number) we cannot say how much time has elapsed for any one of them! Thus, the whole equation will not be known. *For, the whole sum of the radial distances is, to the whole periodic time, as any partial sum of the distances is to its corresponding time.* [Emphasis added]
>
> I consequently began by dividing the eccentric into 360 parts, as if these were least particles, and supposed that within one such part the distance does not change . . . .
>
> However, since this procedure is mechanical and tedious, and since it is impossible to compute the whole equation, given the value for one individual degree [of the eccentric–JT] without the others, I looked around for other means. Considering, that the points of the eccentric are infinite in number, and their radial lines are infinite in number, it struck me, that all the radial lines are contained within the area of the eccentric. I remembered that Archimedes, in seeking the ratio of the cir-

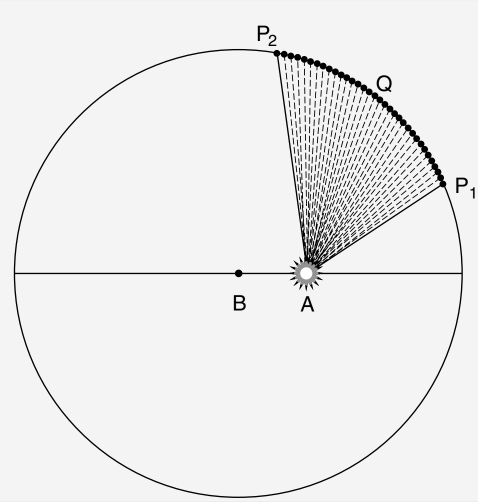


FIGURE 5.3. *Assuming the "momentary" orbital velocities are inversely proportional to the radial distances, Kepler tries to "add up" the radii to determine how much time the planet needs to go from one point of the orbit to another.*

> cumference to the diameter, once divided a circle thus into an infinity of triangles—this being the hidden force of his *reductio ad absurdum.* Accordingly, instead of dividing the circumference, as before, I now cut the *area* of the eccentric into 360 parts, by lines drawn from the point whence the eccentricity is reckoned [*A,* the position of the sun–JT] . . . .

This brief passage marks a crucial breakthrough in the *New Astronomy.* To see more clearly what Kepler has done, on the same diagram as above, mark two positions P_1, P_2 of the planet on the orbit, and draw the radial lines from the sun to those positions—i.e., AP_1 and AP_2. **(Figure 5.3)** Kepler has dropped the idea of using the *length* of the arc between P_1 and P_2 as the appropriate *measure* of the action generating the orbital motion, and turned instead to the *area* of the curvilinear triangle bounded by AP_1, AP_2 and the orbital arc from P_1 to P_2.

We shall later refer to such areas as "orbital sectors." Kepler describes that area as the "sum" of the "infinite number" of radial lines *AQ,* of varying lengths, obtained as *Q* passes through all the positions of the planet from P_1 to P_2! Does he mean this literally? Or, is he not expressing, in metaphorical terms, the *coherence* between the macroscopic process, from P_1 to P_2, and the peculiar "curvature," which governs events within any arbitrarily small interval of that process?

The result, in any case, is a geometrical principle, which Kepler subsequently demonstrated to be empirically valid for the motion of all known planets in their orbits: *The time, which a planet takes in passing from any position P_1 to another position P_2 in its orbit, is proportional*

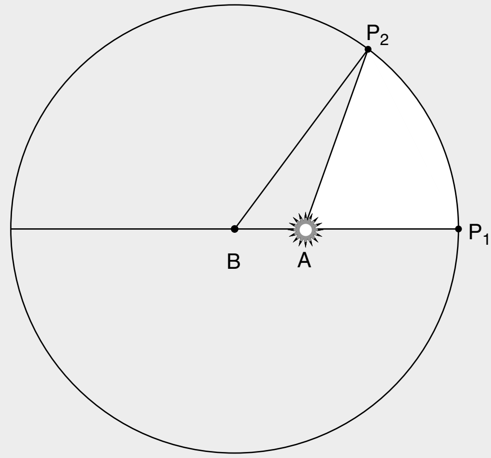


FIGURE 5.4. *Kepler's method for calculating the area swept out by the radial line from the sun to a planet on the assumption that the orbit is an eccentric circle, i.e., a circle whose center B is displaced from the position of the sun A.*

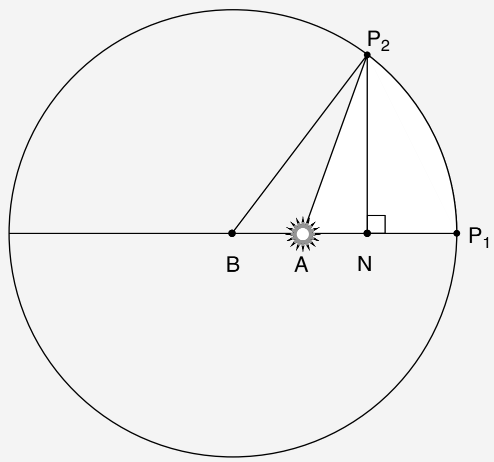
to the area of the sector bounded by the radial lines AP_1, AP_2, and the orbital trajectory P_1P_2, or, in other words, the area swept out by the radial line AP. This is Kepler's famous "Second Law," otherwise known as the "Area Law." All that is needed in addition, to arrive at an extremely precise construction of planetary motion, is to replace the "eccentric circle" approximation, by a true ellipse, as Kepler himself does in the later sections of the *New Astronomy.* We shall attend to that in the next chapter.

Time Produced by Orbital Action?

Are you not struck by something paradoxical in Kepler's formulation? Does he not express himself as if nearly to say, that time is *produced* by the orbital action? Or, does this only seem paradoxical to us (but not to Kepler!), because we have been indoctrinated by the kinematic conceptions of Sarpi, Descartes, and Newton?

There is another paradox implicit here, which Kepler himself emphasized. Sticking for a moment to the eccentric-circle approximation for the orbit, Kepler found a very simple way to calculate the areas of the sectors. In our earlier drawing, choose P_1 to be the intersection of the circumference and the line of apsides passing through *B* and *A*. **(Figure 5.4)** P_1 now represents the position of the planet at the point of perihelion. Take P_2 to be any point on the circumference in the upper half of the circle. If *A* and *B* were at the same place (i.e., if the sun were at the geometrical center of the orbit), then the sectoral area between AP_1 and AP_2 would simply be proportional to the angle formed at *A* between those two lines. Otherwise, we can transform

FIGURE 5.5. *The swept-out area, AP_1P_2, is equal to the circular sector P_1BP_2, minus the triangular area AP_2B.*



the sector in question into a simple, center-based circular sector, by *adding* to it the triangular area ABP_2.

Indeed, as can be seen in **Figure 5.5**, the sum of the two areas is the circular sector between BP_1 and BP_2. The area of the circular sector, on the other hand, is proportional to the angle formed by the radial lines BP_1, BP_2 at the circle's center B, as well as to the circular arc from P_1 to P_2. Turning this around, we can express the sector AP_1P_2, which, according to Kepler, tells us the time elapsed between the two positions, as the result of subtracting the triangle ABP_2 from the sector BP_1P_2. In other words: The time T to go from P_1 to P_2, is proportional to the area AP_1P_2, which in turn is equal to the area of the circular sector between BP_1 and BP_2 minus the area of triangle ABP_2. Of these two areas, the first is proportional to the angle P_1BP_2 at the circle's center and to the circular arc P_1P_2; while the second is equal to the product of the base of triangle ABP_2, namely the length AB, times its height. The height is the length of the perpendicular line P_2N drawn from the orbital position P_2 to the line of apsides, which (up to a factor of the radius) is just the *sine* of the angle P_1BP_2. In this way—leaving aside, for the moment, a certain modification required by the non-circularity of the orbit—Kepler was able to calculate the elapsed times between any two positions in an orbit.

These simple relationships, which are much easier to express in geometrical drawings than in words, are crucial to the whole development up to Gauss. They involve the following peculiarity, highlighted by Kepler: The elapsed time is shown to be a combined function of the indicated *angle* or *circular arc* on the one side, and the length of the perpendicular straight line drawn from P_2 to the line of apsides, on the other. Now, as Kepler notes, *in implicit reference to Nicolaus of Cusa,* those two magnitudes are "heterogeneous"; one is essentially a curved magnitude, the other a straight, linear one. (That is, they are incommensurable; in fact, as Cusa discovered, the curve is "transcendental" to the straight line.) That heterogeneity seems to block our way, when we try to invert Kepler's solution, and to determine the position of a planet after any given elapsed time (i.e., rather than determining the time as it relates to any position). In fact, this is one of the problems which Gauss addressed with his "higher transcendents," including the hypergeometric function.

Let us end this discussion with Kepler's own challenge to the geometers. For the present purposes—deferring some further "dimensionalities" of the problem until Chapter 6—you can read Kepler's technical terms in the following quote in the following way. What Kepler calls the "mean anomaly," is essentially the elapsed time; the term, "eccentric anomaly," refers to the angle subtended by the planetary positions P_1, P_2 as seen from the center B of the circle—i.e., the angle P_1BP_2. Here is Kepler:

> But given the mean anomaly, there is no geometrical method of proceeding to the eccentric anomaly. For, the mean anomaly is composed of two areas, a sector and a triangle. And while the former is measured by the arc of the eccentric, the latter is measured by the sine of that arc. . . . And the ratios between the arcs and their sines are infinite in number [i.e., they are incommensurable as functional "species"–ed.]. So, when we begin with the sum of the two, we cannot say how great the arc is, and how great its sine, corresponding to the sum, unless we were previously to investigate the area resulting from a given arc; that is, unless you were to have constructed tables and to have worked from them subsequently.
>
> That is my opinion. And insofar as it is seen to lack geometrical beauty, I exhort the geometers to solve me this problem:
>
> Given the area of a part of a semicircle and a point on the diameter, to find the arc and the angle at that point, the sides of which angle, and which arc, encloses the given area. Or, to cut the area of a semicircle in a given ratio from any given point on the diameter.
>
> It is enough for me to believe that I could not solve this, *a priori*, owing to the heterogeneity of the arc and sine. Anyone who shows me my error and points the way will be for me the great Apollonius.*

—*JT*

---

* Apollonius of Perga (c. 262-200 B.C.), Greek geometer, author of *On Conic Sections,* the definitive Classical treatise. Drawn by the reputation of the astronomer Aristarchus of Samos, he lived and worked at Alexandria, the great center of learning of the Hellenistic world, where he studied under the successors of Euclid. SEE article, page 100, this issue.–Ed.